Protein engineering of the high-alkaline serine protease PB92 from Bacillus alcalophilus

Teplyakov, A.V.; Lammers, A.A.; Kelders, H.; Kalk, K.H.; Misset, O.; Mulleners, L.J.S.M.; Dijkstra, B.W.

Published in:
Protein Engineering

DOI:
10.1093/protein/5.5.413

Protein Engineering vol.5 no.5 pp.413–420, 1992

Protein engineering of the high-alkaline serine protease PB92 from *Bacillus alcalophilus*: functional and structural consequences of mutation at the S4 substrate binding pocket

A.V.Teplyakov, J.M.van der Laan[1], A.A.Lammers, H.Kelders, K.H.Kalk, O.Misset[1], L.J.S.M.Mulleners[1] and B.W.Dijkstra[2]

BIOSON Research Institute, University of Groningen, Nijenborgh 4, 9747 AG Groningen and [1]Royal Gist-brocades NV, Research and Development, PO Box 1, 2600 MA Delft, The Netherlands

[2]To whom correspondence should be addressed

Serine endoproteases such as trypsins and subtilisins are known to have an extended substrate binding region that interacts with residues P6 to P3′ of a substrate. In order to investigate the structural and functional effects of replacing residues at the S4 substrate binding pocket, the serine protease from the alkalophilic *Bacillus* strain PB92, which shows homology with the subtilisins, was mutated at positions 102 and 126–128. Substitution of Val102 by Trp results in a 12-fold increase in activity towards succinyl-L-Ala-L-Ala-L-Pro-L-Phe-*p*-nitroanilide (sAAPFpNA). An X-ray structure analysis of the V102W mutant shows that the Trp side chain occupies a hydrophobic pocket at the surface of the molecule leaving a narrow crevice for the P4 residue of a substrate. Better binding of sAAPFpNA by the mutant compared with the wild type protein as indicated by the kinetic data might be due to the hydrophobic interaction of Ala P4 of the substrate with the introduced Trp102 side chain. The observed difference in binding of sAAPFpNA by protease PB92 and thermitase, both of which possess a Trp at position 102, is probably related to the amino acid substitutions at positions 105 and 126 (in the protease PB92 numbering). Kinetic data for the variants obtained by random mutation of residues Ser126, Pro127 and Ser128 reveal that the activity towards sAAPFpNA increases when a hydrophobic residue is introduced at position 126. An X-ray diffraction analysis was carried out for the three protease PB92 mutants which have residues Ser126-Pro127-Ser128 replaced by Met-Ala-Gly ('MAG' mutant), Phe-Gln-Ser ('FQS' mutant) and Asn-Ser-Ala ('NSA' mutant). Met126 and Phe126 in the crystal structures of the corresponding mutants are fixed in the same hydrophobic environment as Trp102 in the V102W mutant. In contrast, Asn126 in the 'NSA' mutant is completely disordered in both crystal forms for which the structure has been determined. According to our kinetic measurements none of the mutants with Met, Phe, Leu or Val at position 126 binds sAAPFpNA better than the wild type enzyme. Results of the site-directed mutagenesis at position 127 imply that possible interaction of this residue with a substrate has almost no effect on activity towards sAAPFpNA and casein.

Key words: alkaline protease/*Bacillus alcalophilus*/serine protease/X-ray structure

Introduction

Serine proteases of the subtilisin family (Markland and Smith, 1971) have been extensively studied both to provide insight into the mechanism and specificity of enzyme catalysis (Kraut, 1977) and because of their considerable industrial importance as a protein degrading component of washing powders (Shaw, 1987).

Up to date three-dimensional structures have been determined at high resolution for subtilisins isolated from *Bacillus amyloliquefaciens* [subtilisin BPN′ (SBT)] (McPhalen *et al.*, 1985a; Bott *et al.*, 1988), *Bacillus licheniformis* [subtilisin Carlsberg (SBC)] (McPhalen *et al.*, 1985b; Bode *et al.*, 1987), *Thermoactinomyces vulgaris* [thermitase (TRM)] (Gros *et al.*, 1989; Teplyakov *et al.*, 1990) and *Tritirachium album* (proteinase K) (Betzel *et al.*, 1988). The crystal structures served as a basis for a number of protein engineering studies which resulted in mutants with altered stability and kinetic properties (Bryan *et al.*, 1986; Estell *et al.*, 1986; Russell and Fersht, 1987; Wells *et al.*, 1987; Pantoliano *et al.*, 1988, 1989; Carter *et al.*, 1990).

Substrate specificity of subtilisins is broad. The observation that the hydrolysis rate of esters of N-terminally blocked amino acids was dependent on the type of amino acid suggested that the specificity is determined by the interaction with the P1 residue of a peptide substrate (Glazer, 1967; Morihara and Tsuzuki, 1969) [residues of a substrate and the corresponding binding sites of the enzyme are named according to the Schechter and Berger (1967) notation]. Morihara *et al.* (1970) showed that the specificity is also affected by at least five amino acid residues (P4–P2′) surrounding the sensitive residue P1. X-ray structure analysis of various subtilisin–inhibitor complexes (Hirono *et al.*, 1984; Bode *et al.*, 1987; McPhalen and James, 1988; Gros *et al.*, 1989) provided direct evidence for the existence of the extended substrate binding region that interacts with residues P6–P3′ of a substrate. Moreover it was pointed out that the interaction involving P4 is almost as significant as the main P1–S1 interaction (Robertus *et al.*, 1972; Hirono *et al.*, 1984). On the other hand, mutation of the P4 residue of the *Streptomyces* subtilisin inhibitor, SSI, in which Met was replaced by Gly, Ala and Phe (Kojima *et al.*, 1990) appeared to have hardly any effect on its inhibitory activity towards SBT. To investigate further the importance of the interaction at the S4 site we decided to mutate residues forming this site in the serine protease isolated from the *Bacillus alcalophilus* strain PB92 (SBA) (van der Laan *et al.*, 1991).

From the crystal structures of subtilisin–inhibitor complexes (Hirono *et al.*, 1984; Bode *et al.*, 1987; McPhalen and James, 1988; Gros *et al.*, 1989) it is known that the S4 site is formed by residues from two loops at the surface of the molecule. These loops contain residues 100–105 and 124–129 in SBA (102–107 and 126–131 in the SBT sequence). Flexibility of these loops is believed to be an important feature of the induced-fit mechanism of substrate binding (McPhalen and James, 1988; Gros *et al.*, 1989). According to Bode *et al.* (1987) and Gros *et al.* (1989) the P4 residue makes a number of contacts in the relatively hydrophobic P4 pocket of the enzyme. An aromatic residue at the position equivalent to 102 in SBA (Tyr104 in SBT, Tyr103 in SBC, Trp112 in TRM) seems to contribute substantially to the formation of this pocket. In SBA the corresponding position is occupied by Val102 which we mutated to Trp in the present study. Residues from the other loop involved

in the formation of the S4 binding site, namely residues 126–128 of SBA, were subjected to a random mutagenesis and the variants obtained were studied with respect to their kinetic properties and three-dimensional structures. As these residues are located in a surface loop and at a distance from the active centre we expected no significant conformational changes in the core of the protein and no direct influence on the catalytic mechanism.

Materials and methods

Protein

High alkaline serine protease used for the experiments was isolated from the *Bacillus* strain PB92 (van Eekelen *et al.*, 1989; van der Laan *et al.*, 1991). The polypeptide chain of the protein consists of 269 amino acid residues which show a reasonable degree of homology with subtilisins from *Bacillus* species (61% of identical residues compared with SBC). The crystal structure of the wild type enzyme has been determined at 1.75 Å resolution (van der Laan *et al.*, 1992; see accompanying article).

Mutation procedures and DNA construction

The gene of the serine protease from the *Bacillus* strain PB92 (ATCC 31408) has been cloned into the *Bacillus subtilis* strain 1-A40 (Young *et al.*, 1969). For mutagenesis the gene was recloned into phage M13mp11 (Messing *et al.*, 1981) and transfected to *Escherichia coli* JM101 (Cohen *et al.*, 1972). After phage propagation in *E.coli* JM101 dsDNA was isolated according to the method described by Birnboim and Doly (1979). The insert and its orientation were checked by DNA sequencing (Sanger *et al.*, 1977).

The method of mutagenesis was based on the gapped-duplex approach (Kramer *et al.*, 1984) and a phage–plasmid hybrid. Instead of *E.coli* WK30-3, *E.coli* JM101 was used for the selection of the mutants. M13mp19 dsDNA was digested with *Eco*RI and *Hin*dIII and subsequently annealed with the vector M13 M1 which contained the protease gene. An oligonucleotide with the desired mutation was annealed into the gap which was closed with polymerase and T4 ligase. The 22-base oligonucleotides were used to create the specific mutations. Region-specific mutation used to create more mutations at a time in a specific DNA sequence was performed using an oligonucleotide preparation with a length of 40 nucleotides. Nucleotides were randomly incorporated in the sites corresponding to the amino acid(s) to be mutated. Potential mutants were checked by sequence analysis. The entire single strand gap was sequenced to check for the absence of secondary mutations. To produce the mutant the DNA fragment of M13 M1 containing the desired mutations was subcloned into a pM58 plasmid which contained a neomycin resistance marker but lacked the fragment with the desired mutations. Both digests were ligated and transformed to *B.subtilis* DB104 using the method of Anagnostopoulos and Spizizen (1961) and selected for neomycin resistance and protease activity. DNA of protease producing transformants was isolated according to the method described by Birnboim and Doly (1979) and was characterized by restriction enzyme analysis.

Production and purification of SBA mutants

Transformants of DB104 which were determined to contain the vector with the mutated protease gene were inoculated in 10 ml Tryptone Soya Broth (TSB) with 20 μg/ml neomycin and were incubated for 24 h at 37°C. Samples of the culture were inoculated in 500 ml shake flasks containing 100 ml protease production medium, 12.5 g/l yeast extract (Difco), 0.97 g/l $CaCl_2 \cdot 6H_2O$, 2.25 g/l $MgCl_2 \cdot 6H_2O$, 20 mg/ml $MnSO_4 \cdot 4H_2O$, 1 mg/l $CoCl_2 \cdot 6H_2O$, 0.5 g/l citrate, 0.5 ml/l antifoam 5693, 6% (w/v) maltose, 0.2 M phosphate buffer, pH 6.8, and 20 μg/ml neomycin.

The wild type protease and the mutants were purified by cation exchange chromatography using S-Sepharose (Pharmacia). Fractions containing protease activity were pooled and concentrated by ultrafiltration using an Amicon stirred cell. The purity of the proteases was checked by SDS–PAGE and by high performance liquid chromatography (HPLC).

Activity measurements

The protease concentration was determined via extinction measurements at 280 nm using a molar extinction coefficient of 26 100 M^{-1} cm^{-1} which was calculated from the number of tryptophans and tyrosines per enzyme molecule. For the mutants with an altered number of tryptophans or tyrosines the corresponding corrections were introduced. An estimation of the number of active enzyme molecules was obtained by active site titrations which were performed with phenylmethylsulfonylfluoride (PMSF).

Activity was measured on succinyl-L-alanyl-L-alanyl-L-prolyl-L-phenylalanyl-paranitroanilide (sAAPFpNA) and casein. Concentration of yellow paranitroanilide (pNA) was measured spectrophotometrically at 410 nm; ϵ_M = 8480 M^{-1} cm^{-1} (DelMar *et al.*, 1979). The kinetic parameters k_{cat} and K_m were obtained from initial rate measurements at various substrate concentrations (from 0.1 to 6.0 mM) at 25°C in 0.1 M Tris–HCl, pH 8.6, with 0.1 M NaCl. Activity on casein at pH 10.0 was quantified in terms of Alkaline Delft Units (ADUs) (van Velzen, 1970) according to the method which was based on the casein assays described by Kunitz (1947) and Laskowski (1955). Thus, one ADU may be defined as the amount of an enzyme which after incubation at 40°C for 40 min at 6 cm^3 of a standard casein solution and after precipitation by a further 5 cm^3 of a standard trichloroacetic acid solution will give an absorbance difference in the supernatant of 0.022 at 275 nm relative to an appropriate blank.

X-ray analysis

An X-ray diffraction analysis was carried out for the three SBA mutants having the following tripeptides in position 126–128: Phe-Gln-Ser (hereafter 'FQS' mutant), Met-Ala-Gly ('MAG' mutant), Asn-Ser-Ala ('NSA' mutant) and for the single-site V102W mutant. Crystals of the mutants were obtained by the vapour diffusion method in hanging drops from 30% PEG 6000 solution in 0.1 M potassium phosphate buffer, pH 6.0. All crystals belong to space group $P2_12_12_1$ but have different unit cell parameters (Table I). The crystal structure of the 'NSA' mutant was determined in two crystal forms. X-ray data were collected on a FAST diffractometer with a TV area detector (Enraf-Nonius, Delft). One crystal of each mutant was used for data collection. Parameters of the X-ray experiments are presented in Table II.

The structures of the mutants were determined by molecular replacement using the atomic model of wild type SBA (van der

Table I. Crystal forms of the SBA mutants

	NSA-I	NSA-II	FQS	MAG	V102W
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
a, Å	47.7	40.2	39.8	53.3	53.1
b, Å	62.3	62.2	62.3	61.5	61.5
c, Å	75.7	92.1	92.0	75.3	74.9
Crystal form	I	II	II	III	III

Laan *et al.*, 1992). All structures were refined by the TNT restrained least-squares procedure (Tronrud *et al.*, 1987). Cycles of positional refinement were followed by an individual atomic temperature factor refinement. Solvent molecules were added to the models on the basis of (F_o-F_c) difference electron density maps calculated with σ_A weights (Read, 1986). Minor manual corrections of the models using the FRODO program (Jones, 1978) on an Evans and Sutherland PS390 were necessary to adjust the amino acid sequences.

Results and discussion

V102W mutant

Specific activities on sAAPFpNA measured for wild type SBA and the V102W mutant as well as for SBT and TRM are given in Table III. Substitution of Val102 by Trp in SBA causes a 12-fold increase in activity towards sAAPFpNA. There is an obvious shift of the kinetic parameters K_m and k_{cat} upon mutation towards those of TRM which also has Trp at the corresponding position.

To analyse the atomic details of the mutation, the crystal structure of the V102W mutant was determined by X-ray diffraction methods at 1.8 Å resolution. The atomic model has been refined to a crystallographic *R*-factor ($\Sigma|F_o-F_c|/\Sigma F_o$, where F_o and F_c are observed and calculated structure amplitudes, respectively) of 13.9%. The estimated root-mean-square (r.m.s.) error in the atomic coordinates derived from the σ_A plot of Read (1986) is 0.12 Å. The refinement statistics are presented in Table IV.

Superposition of this model on the wild type structure gave r.m.s. differences of 0.32 Å for the main chain atoms and 0.56 Å for all 1879 common atoms (Table IV). Both values are somewhat larger than the estimated r.m.s. error mainly because of rearrangement of a few residues on the surface of the molecule due to different packing in crystal forms I (wild type) and III (V102W). The superposition of the V102W mutant on the structure of the 'MAG' mutant which was determined in the same crystal form III (see below) resulted in r.m.s. differences of 0.11 Å for the main chain and 0.23 Å for all atoms. These values agree with the finding that mutation of residue 102 in SBA causes no significant conformational changes.

Substitution of Val102 by Trp results in the formation of a relatively closed P4 pocket (Figure 1). Modelling of sAAPFpNA binding by SBA based on the TRM–eglin c structure (Gros *et al.*, 1989) indicates that the side chain of Trp102 would be at van der Waals distance from the expected position of residue P4 (3.6 Å between C^{β} of P4 and $C^{\eta 2}$ of Trp102) while Val at position 102 would be more than 5 Å away from P4. All other residues at the S4 site of SBA keep their positions upon the mutation. We suggest that better binding of sAAPFpNA by the mutant compared with the wild type protein as indicated by the kinetic data (Table III) might be due to the additional interaction of Ala P4 of the substrate with the Trp102 side chain. Because the shift in backbone position of the 124–129 loop (Figure 1) is observed without exception in all structures which were solved for crystal form III, even if they do not contain any mutation at positions 102 or 126–129, we believe that the observed shift is due to crystal packing. It suggests that the 124–129 loop might be rather mobile in solution. Upon crystallization its conformation is fixed by crystal contacts.

Table II. X-ray diffraction data for the SBA mutants

	NSA-I	NSA-II	FQS	MAG	V102W
Resolution, Å	1.7	2.0	1.9	1.7	1.8
Total number of reflections	32072	28186	47507	47850	51400
$R_{merge} = \Sigma\|I_i-<I>\|/\Sigma I_i$, %	5.2	4.4	5.7	4.1	4.5
Number of unique reflections	19966	14437	16972	22790	21700
Completeness, %					
total	78.4	89.0	90.3	81.6	92.9
last shell[a]	38.9	58.1	83.5	54.5	82.3

[a]The highest resolution shell contains 1000 theoretically expected reflections.

Table III. Kinetic parameters for sAAPFpNA hydrolysis by subtilisins

Enzyme	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})
SBA, wild type	1.20	103	90
SBA, V102W	0.40	447	1120
SBT	0.15	35	230
TRM	0.03	320	10700

Table IV. Refinement statistics for the SBA mutants

Mutants	NSA-I	NSA-II	FQS	MAG	V102W
Resolution, Å	7.0–1.75	7.0–2.0	7.0–1.9	7.0–1.75	7.0–1.8
$R = \Sigma\|F_o-F_c\|/\Sigma F_o$[a] (%)	14.9	15.3	15.3	14.9	13.9
Number of atoms:					
protein	1883	1883	1892	1882	1889
water	174	151	178	218	195
cations (Ca^{2+}/K^+)	2	2	2	1	1
R.m.s. deviation from ideality:					
bond lengths (Å)	0.014	0.014	0.008	0.014	0.015
bond angles, degrees	2.8	3.3	2.6	2.7	2.9
planar groups (Å)	0.017	0.013	0.009	0.015	0.016
non-bonded contacts (Å)	0.050	0.044	0.045	0.059	0.066
R.m.s. deviation in B-factors of bonded atoms ($Å^2$)	2.0	2.0	1.4	1.8	2.1
Estimated r.m.s. error[b] (Å)	0.07	0.13	0.12	0.09	0.12
R.m.s. deviation from wild type:					
main chain atoms (Å)	0.12	0.27	0.32	0.30	0.32
all atoms (Å)	0.17	0.39	0.57	0.56	0.56

[a]All measured reflections in the defined resolution range are included in the R-factor calculation.
[b]Root-mean-square error in atomic coordinates estimated from the σ_A plot of Read (1986).

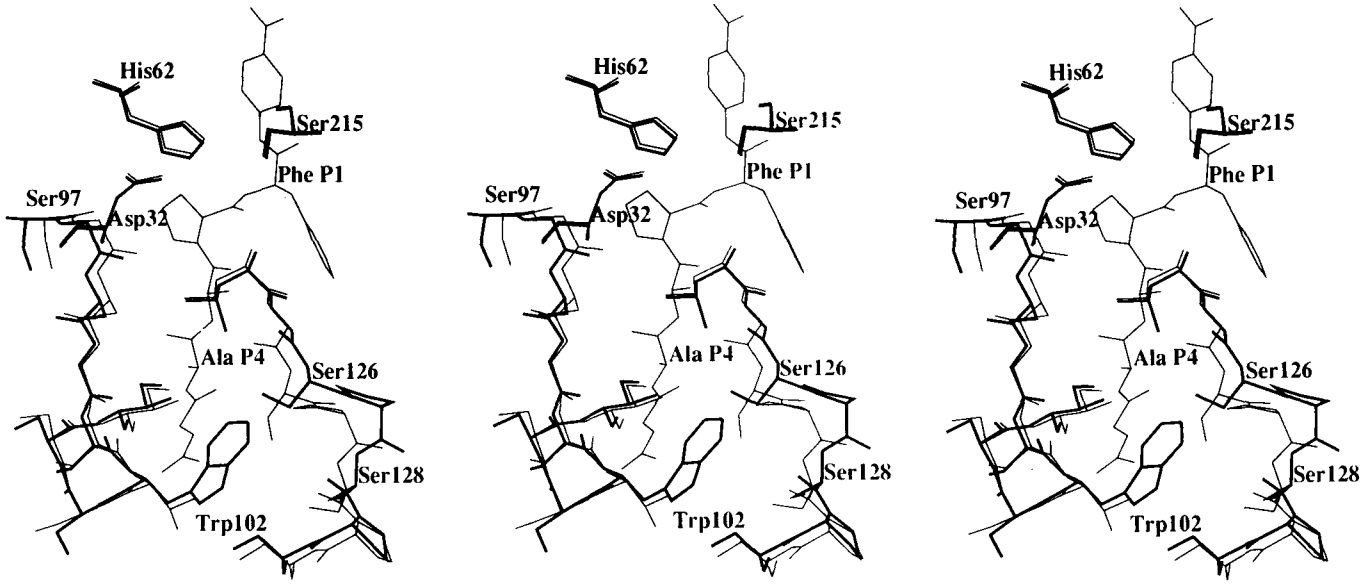

Fig. 1. The S4 binding sites of the V102W mutant (in thick lines) and wild type SBA (in thin lines) after an overall superposition of the structures. The substrate sAAPFpNA labelled at residues P1 and P4 is modelled on the basis of the TRM–eglin c structure (Gros *et al.*, 1989). The P4 to P1 residues in eglin c are Pro-Ser-Thr-Leu.

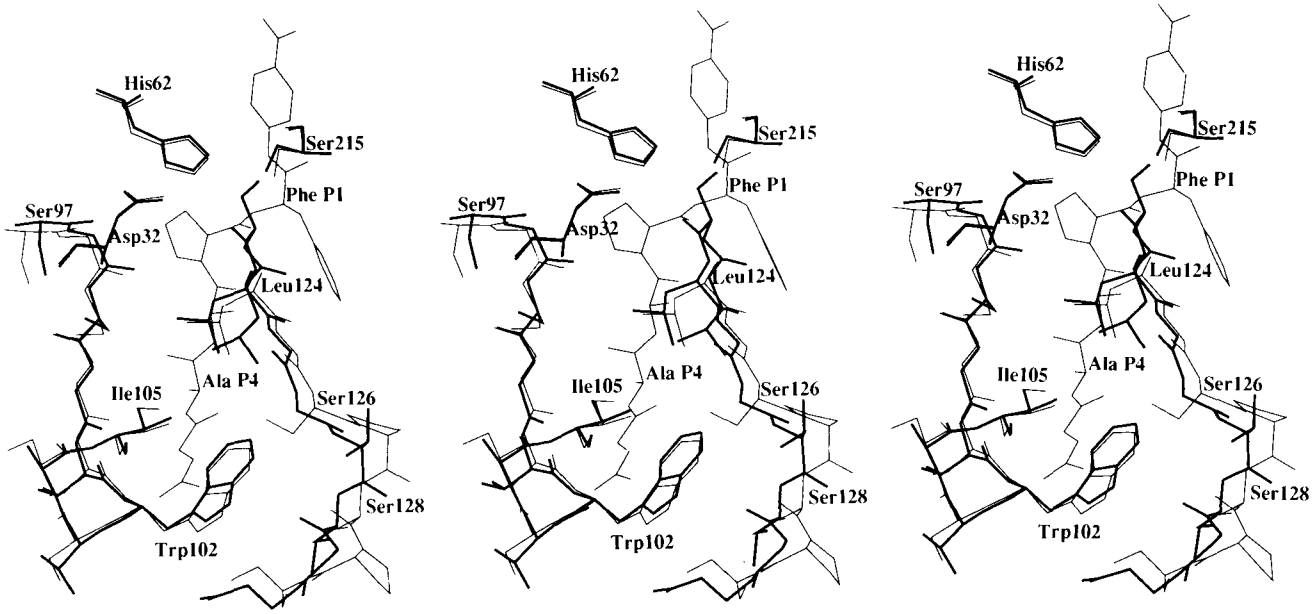

Fig. 2. The S4 binding sites of the V102W mutant of SBA (in thin lines) and TRM (in thick lines) after an overall superposition of the structures. The substrate sAAPFpNA labelled at residues P1 and P4 is modelled on the basis of the TRM–eglin c structure (Gros *et al.*, 1989).

Whereas introduction of Trp102 in SBA leads to a better binding of sAAPFpNA there is still a difference in K_m of one order of magnitude with respect to TRM which also possesses Trp at this position. The side chain of Trp102 in the mutant has the same conformation as in TRM as can be judged from a comparison of the side chain dihedral angles: ($\chi 1$, $\chi 2$) = (−66°, 125°) in SBA and (−79°, 131°) in TRM. When the two structures are superimposed using 1048 common main chain atoms (with an r.m.s. deviation of 1.22 Å) the Trp residues overlap with an r.m.s. deviation of only 0.50 Å (Figure 2). This similarity suggests that the residue at position 102 determines the binding properties only to a minor extent and that there should be another reason for the observed difference in binding of sAAPFpNA by SBA and TRM.

Inspection of the binding sites from P4 to P2′ in both proteins (Figure 2) revealed for residues in close contact with the substrate only a few amino acid substitutions: at positions 105, 126, 160 and 212, while all the other residues are conserved. The amino acid difference at position 212 does not directly affect substrate binding because its side chain points away from the binding site as was noted by Gros *et al.* (1989). Residue 106 is located at the S1 site and can interact with the side chain of residue P1 only if it is long enough. In the TRM–eglin c complex (Gros *et al.*, 1989) the shortest distance between Asn170, which corresponds

to Ser160 in SBA and Leu37 of eglin c (residue P1) is that between $O^{\delta 1}$ of Asn170 and $C^{\delta 2}$ of Leu37 (4.58 Å). The other two residues, 105 and 126 in SBA, are located at the S4 site and are of apparent importance for substrate binding. Ser126 in SBA instead of Gly in TRM is in very close contact with the side chain of Trp102: 3.17 Å between Ser126 C^{β} and Trp102 $C^{\zeta 2}$. This steric hindrance can restrict a possible induced-fit movement of the loop 124–130 in SBA when binding a substrate. The other residue that might be responsible for the observed difference (or at least in part) in binding constants between SBA and TRM is Ile105 (Val115 in TRM). The side chain of this residue forms the bottom of the relatively deep P4 pocket (Figure 2). In the TRM–eglin c complex (Gros *et al.*, 1989) Val115 $C^{\gamma 1}$ is at a distance of 4.1 Å from C^{β} of residue P4 (Pro) of the inhibitor. The corresponding distance in SBA (Ile105 $C^{\delta 1}$ to P4 C^{β}) derived from the superposition of the models is 1 Å shorter. Clearly the additional methyl group of Ile can impose serious steric restrictions on binding of non-glycine residues at the S4 site.

Mutants at position 126–128

Residues 126–128 form the central part of a flexible loop at the surface of the molecule. Judging from the crystal structures of several subtilisin–inhibitor complexes (Hirono *et al.*, 1984; Bode *et al.*, 1987; McPhalen and James, 1988; Gros *et al.*, 1989) these residues can interact with residues P4–P6 of a substrate. The tripeptide 126–128 in SBA was subjected to a random mutagenesis and the variants obtained were tested on plates with respect to their activity on casein. For a number of the mutants which retained at least 50% of the activity of the wild type enzyme the nucleotide sequences were determined and also the kinetic parameters of sAAPFpNA hydrolysis were measured. The results are summarized in Table V.

Sequencing of the mutants revealed a replacement of Gly116 by Val which accidentally occurred upon cloning of the gene. However, this residue is located on the opposite side of the molecule with respect to the substrate binding cleft. The X-ray structures which contain this additional mutation show that the replacement does not lead to any conformational changes. Therefore this substitution is believed to have no effect on the binding of substrates.

From inspection of Table V one can see that the mutations do influence the kinetic parameters of SBA for sAAPFpNA hydrolysis. As the substrate is short and lacks residues P5 and P6 we can suggest that the differences are most probably related to the interaction of the P4 residue of the substrate with residues forming the S4 site of the enzyme, particularly with residue 126. The second observation is that high turnover numbers (k_{cat} >200 s^{-1}) correspond to the mutants with hydrophobic residues (Met, Leu, Val, Phe) at position 126 whereas the mutants with other residues at this position (Ser, Gly, His, Asn) show significantly lower activity on sAAPFpNA.

To investigate the influence of residue 127 on activity towards both substrates, sAAPFpNA and casein, we obtained two series of mutants of type P127X: one with Ser and another with Phe at position 126. The results of this study are presented in Table VI. As expected there is almost no effect on kinetic parameters measured for a short substrate, but there is also very little effect on activity towards casein for both series of the mutants. The latter indicates that possible interaction between residue 127 and a substrate does not influence activity of the enzyme.

In order to explain the properties of the mutants in terms of their three-dimensional structures we carried out an X-ray diffraction analysis of three mutants with sequences Phe-Gln-Ser (FQS), Met-Ala-Gly (MAG) and Asn-Ser-Ala (NSA) in position 126–128. The refinement statistics for these mutants are given in Table IV. The three-dimensional structures of the 'FQS', 'MAG' and 'NSA' mutants of SBA are highly similar. Superposition of these structures on the structure of the wild type protein gave r.m.s. differences of ~0.3 Å for the main chain atoms. For the first crystal form of the 'NSA' mutant the difference is much smaller than for the other structures including the second crystal form of the 'NSA' mutant. This is a consequence of the identical packing of the native and 'NSA' molecules in the same crystal form. For the other structures the relatively large r.m.s. differences of ~0.6 Å calculated for all common atoms are due to a few residues deviating in atomic positions by >1 Å. These are the side chains of nine residues (Asn42, Gln107, Glu110, Asn114, Gln135, Arg143, Arg164, Lys231, Arg269) and three short segments of the polypeptide chain (52–54, 96–97, 125–127). All these residues lie on the surface of the protein molecule and possess different conformations because of different intermolecular contacts in the three crystal forms.

Table V. Kinetic parameters for sAAPFpNA hydrolysis by SBA variants mutated at position 126–128

SBA variant[a]	K_m (mM)	k_{cat} (s^{-1})
Wild type (SPS)	1.2	103
MAG	2.0	356
HYS	1.8	47
GQI	3.2	41
NSA	2.6	69
LNV	4.1	265
VMS	2.1	207
LQA	4.9	264
FQS	8.1	210
FLT	11.0	200

All variants except MAG and FQS also have mutation G116V (see text).
[a]Residues at position 126–128 are given in the one-letter code.

Table VI. Specific activity towards casein and kinetic parameters for sAAPFpNA hydrolysis by SBA variants mutated at position 127

SBA variant[a]	Casein[b] (%)	K_m (mM)	k_{cat} (s^{-1})
SPS (wild type)	100	1.2	103
SAS	113	1.2	144
SNS	103	1.0	130
STS	92	1.2	140
SHS	82	1.6	165
SKS	78	0.9	159
SES	77	1.6	137
SDS	n.d.[c]	1.3	113
FHS	26	7.8	197
FKS	22	6.3	224
FQS	20	8.1	210
FAS	17	10.3	225
FDS	9	8.2	114
FTS	n.d.[c]	8.1	206

[a]Residues at positions 126–128 are given in the one-letter code.
[b]Wild type SBA has 100% specific activity on casein by definition.
[c]Not determined.

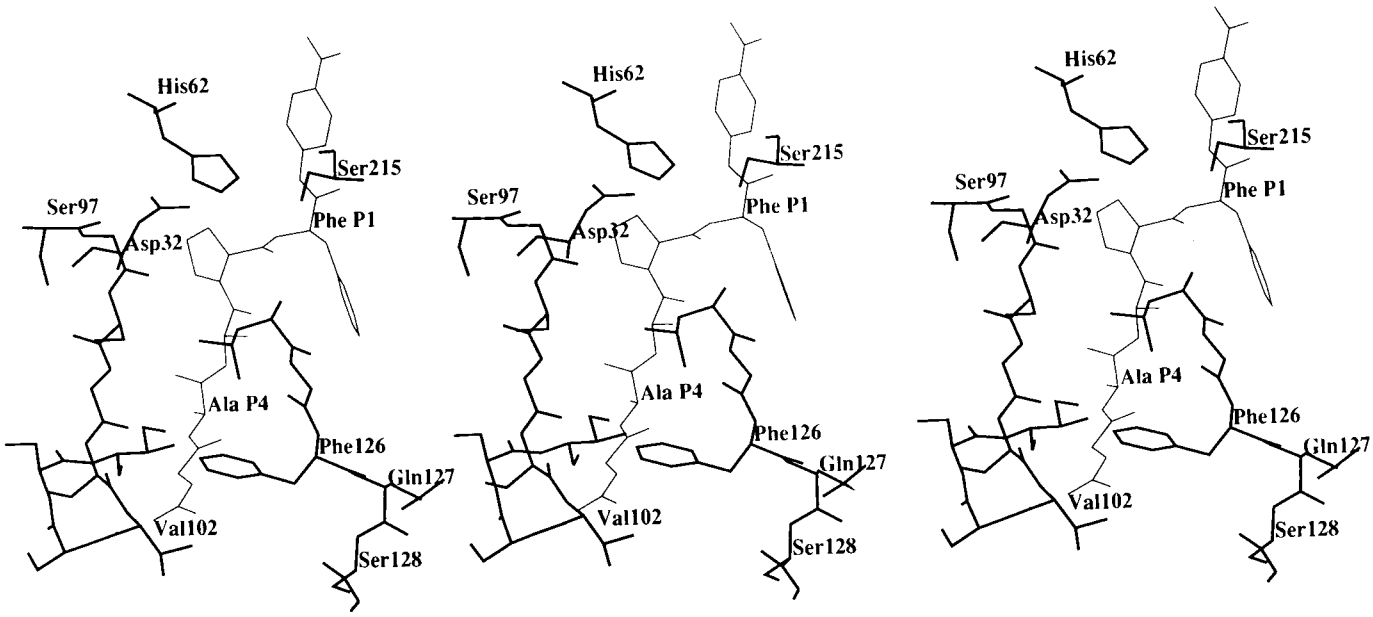

Fig. 3. The S4 binding site of the 'FQS' mutant of SBA. The substrate sAAPFpNA labelled at residues P1 and P4 is modelled on the basis of the TRM–eglin c structure (Gros *et al.*, 1989).

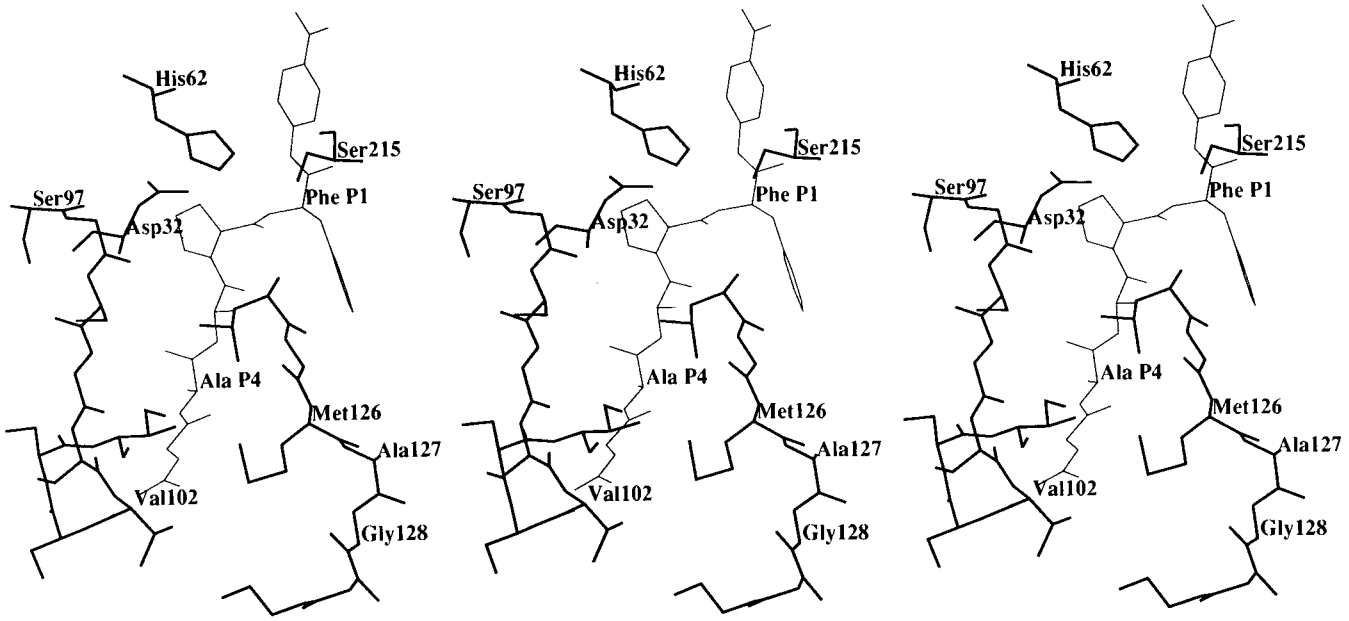

Fig. 4. The S4 binding site of the 'MAG' mutant of SBA. The substrate sAAPFpNA labelled at residues P1 and P4 is modelled on the basis of the TRM–eglin c structure (Gros *et al.*, 1989).

The largest main chain differences occur in surface loops all of which contain glycine residues. The loop 51–54 (in SBA numbering) has different conformations in all subtilisins with known three-dimensional structure (Bode *et al.*, 1987; McPhalen and James, 1988; Betzel *et al.*, 1988; Gros *et al.*, 1989). In TRM its flexibility is limited due to binding of a calcium ion at the so-called 'medium site' (Gros *et al.*, 1989; Teplyakov *et al.*, 1990). There is no such site in the other subtilisins including SBA. Loops 95–98 and 125–128 are involved in substrate binding: the former at the S2–S3 site, the latter at the S4–S6 site. It has been suggested (McPhalen and James, 1988; Gros *et al.*, 1992) that conformational flexibility of these loops is necessary for induced-fit binding of a substrate.

The loop 125–128 contains the three residues which were mutated in the present study. Thus it is very important to separate the effects of crystal packing from the effects of amino acid substitutions on the conformation of this loop. The electron density at residues 126–128 is very clear in the 'FQS' and 'MAG' mutants allowing unambiguous positioning of all atoms in this loop. Both residues at position 126, Phe in the 'FQS' mutant and Met in the 'MAG' mutant, occupy a hydrophobic pocket at the surface of the molecule formed by residues Val102, Ile105, Leu124, Leu133, Pro162 (Figures 3 and 4). In the V102W mutant this pocket was filled by the side chain of Trp102 (Figure 1).

In the 'NSA' mutant in both crystal forms there is lack of

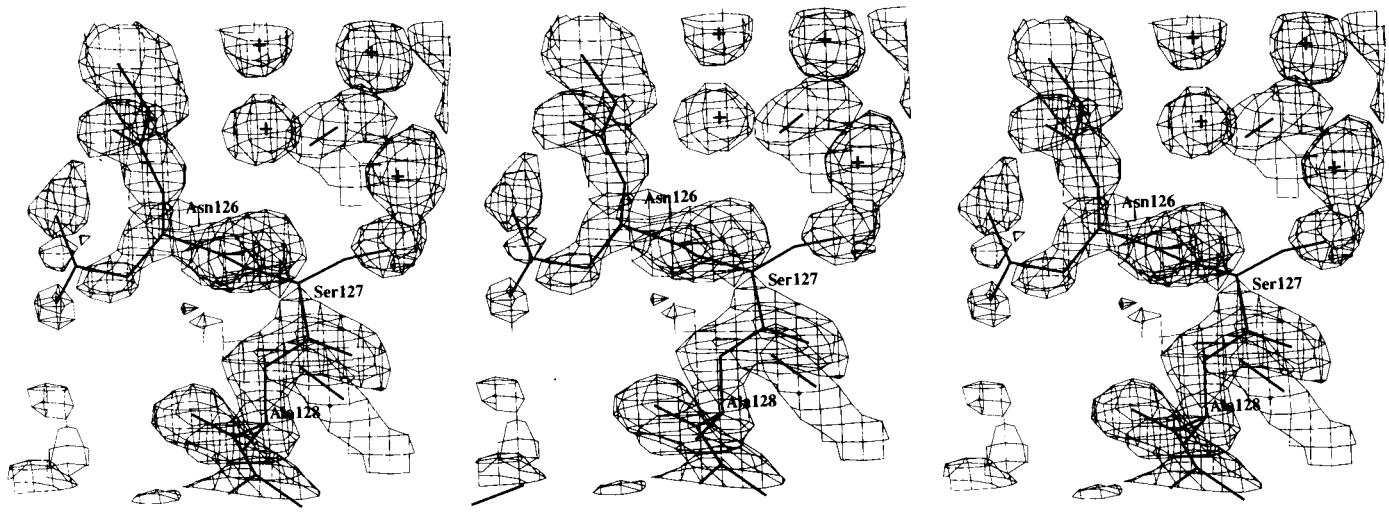

Fig. 5. A fragment of a (F_o-F_c) electron density map and the atomic model at the S4 binding site of the 'NSA-I' mutant of SBA. The map is contoured at a level of one standard deviation. Residues 126–128 were omitted from the model for the phase calculation.

electron density for the side chain of residue 126 (Figure 5). In contrast to Asn126, Phe126 in the 'FQS' mutant in one of these crystal forms (II) is well fixed. Met126 in the 'MAG' mutant is also fixed although in this crystal form (III) the loop 126–128 is not at all involved in contacts with neighbouring molecules and one could expect more conformational freedom for a Met side chain. Thus it seems unlikely that different conformational flexibility of the side chain of residue 126 in the three SBA mutants is caused by the crystal packing effects. Probably the hydrophobic character of the residues surrounding position 126 is not an ideal environment for the Asn side chain which shows as a result a high degree of conformational mobility.

Hydrophobic residues at position 126 (Met, Phe and presumably Leu and Val) are fixed at the S4 site. From the kinetic data presented in Table V one can conclude that high activity towards sAAPFpNA is related to the ability of the enzyme to provide a hydrophobic environment at the S4 site. However, this ability does not correlate with the binding constants, K_m, which are in the same range for all the mutants listed. Moreover there is a significant difference in kinetic parameters between the V102W mutant and the 'hydrophobic' mutants, particularly the S126F series, although both provide the hydrophobic S4 site.

Clearly binding of a substrate is determined by the conformation and properties of individual residues forming the binding site rather than by one overall feature of the binding pocket such as hydrophobicity. From this point of view we could expect a similar kinetic behaviour for the mutants with Leu and Val at position 126, and we do observe it (see Table V), because of similar properties and size of these residues. On the other hand there may be a significant difference between V102W and S126F mutants due to a different location of the introduced aromatic residue in the P4 pocket (compare Figures 1 and 3). In the 'FQS' mutant the side chain of Phe126 comes too close (up to 1.5 Å) to the expected position of the P4 residue and apparently has to move out of the pocket when binding a substrate. This may be the reason for a somewhat higher K_m for all the mutants possessing Phe126 (Tables V and VI).

The results of this study confirm the importance of the interaction at the S4 site for the activity of protease PB92. By changing residues at this site it appears possible to obtain variants with an altered activity towards a certain substrate.

Acknowledgements

The research was sponsored by International Bio-Synthetics (IBIS), Rijswijk, The Netherlands.

References

Anagnostopoulos,C. and Spizizen,J. (1961) *J. Bacteriol.*, **81**, 741–746.
Betzel,Ch., Pal,G.P. and Saenger,W. (1988) *Acta Crystallogr.*, **B44**, 163–172.
Birnboim,H.C. and Doly,J. (1979) *Nucleic Acids Res.*, **7**, 1513–1523.
Bode,W., Papamokos,E. and Musil,D. (1987) *Eur. J. Biochem.*, **166**, 673–692.
Bott,R., Ultsch,M., Kossiakoff,A., Graycar,T., Katz,B. and Powers,D. (1988) *J. Biol. Chem.*, **263**, 7895–7906.
Bryan,P.N., Rollence,M.L., Pantoliano,M.W., Wood,J., Finzel,B.C., Gilliland,G.L., Howard,A.J. and Poulos,T.L. (1986) *Proteins*, **1**, 326–334.
Carter,P., Nilsson,B., Burnier,J.P., Burdick,D. and Wells,J.A. (1990) *Proteins*, **6**, 240–248.
Cohen,S.N., Chang,A.C.Y. and Hsu,L. (1972) *Proc. Natl Acad. Sci. USA*, **69**, 2110–2114.
DelMar,E.G., Largman,C., Brodrick,J.W. and Geokas,M.S. (1979) *Anal. Biochem.*, **99**, 316–320.
van Eekelen,C.A., van der Laan,J.C., Mulleners,L.J.S.M., Misset,O., Cuperus,R.A. and Lensink,J.H.A. (1989) European Patent Application 0328229-A1.
Estell,D.A., Graycar,T.P., Miller,J.V., Powers,D.B., Burnier,J.P., Ng,P.G. and Wells,J.A. (1986) *Science*, **233**, 659–663.
Glazer,A.N. (1967) *J. Biol. Chem.*, **242**, 433–436.
Gros,P., Betzel,Ch., Dauter,Z., Wilson,K.S. and Hol,W.G.J. (1989) *J. Mol. Biol.*, **210**, 347–367.
Gros,P., Teplyakov,A. and Hol,W.G.J. (1992) *Proteins*, **12**, 63–74.
Hirono,S., Akagawa,H., Mitsui,Y. and Iitaka,Y. (1984) *J. Mol. Biol.*, **178**, 389–413.
Jones,J.A. (1978) *J. Appl. Crystallogr.*, **11**, 268–272.
Kojima,S., Kumagai,I. and Miura,K. (1990) *Protein Engng*, **3**, 527–530.
Kramer,W., Drutsa,V., Jansen,H.W., Kramer,B., Plugfelder,M. and Fritz,H.J. (1984) *Nucleic Acids Res.*, **12**, 9441–9456.
Kraut,J. (1977) *Annu. Rev. Biochem.*, **46**, 331–358.
Kunitz,M.J. (1947) *J. Gen. Physiol.*, **30**, 291.
van der Laan,J.C., Gerritse,G., Mulleners,L.J.S.M., van der Hoek,R.A.C. and Quax,W.J. (1991) *Appl. Environ. Microbiol.*, **57**, 901–909.
van der Laan,J.M., Taplyakov,A.V., Kelders,H., Kalk,K.H., Misset,O., Mulleners,L.J.S.M. an Dijkstra,B.W. (1992) *Protein Engng*, **5**, 411–415.
Laskowski,M. (1955) *Methods Enzymol.*, **2**, 26–36.
Markland,F.S. and Smith,E.L. (1971) In Boyer,P.D. (ed.), *The Enzymes*. Academic Press, New York, Vol. 3, pp. 561–608.
McPhalen,C.A. and James,M.N.G. (1988) *Biochemistry*, **27**, 6582–6598.
McPhalen,C.A., Svendsen,I., Jonassen,I. and James,M.N.G. (1985a) *Proc. Natl Acad. Sci. USA*, **82**, 7242–7246.
McPhalen,C.A., Schnebli,H.P. and James,M.N.G. (1985b) *FEBS Lett.*, **188**, 55–58.

Messing,J., Crea,R. and Seeburg,P.H. (1981) *Nucleic Acids Res.*, **9**, 303.
Morihara,K. and Tsuzuki,H. (1969) *Arch. Biochem. Biophys.*, **129**, 620–634.
Morihara,K., Oka,T. and Tsuzuki,H. (1970) *Arch. Biochem. Biophys.*, **138**, 515–525.
Pantoliano,M.W., Whitlow,M., Wood,J.F., Rollence,M.L., Finzel,B.C., Gilliland,G.L., Poulos,T.L. and Bryan,P.N. (1988) *Biochemistry*, **27**, 8311–8317.
Pantoliano,M.W., Whitlow,M., Wood,J., Dodd,S.W., Hardman,K.D., Rollence,M.L. and Bryan,P.N. (1989) *Biochemistry*, **28**, 7205–7213.
Read,R.J. (1986) *Acta Crystallogr.*, **A42**, 140–149.
Robertus,J.D., Alden,R.A., Birktoft,J.J., Kraut,J., Powers,J.C. and Wilcox,P.E. (1972) *Biochemistry*, **11**, 2439–2449.
Russell,A.J. and Fersht,A.R. (1987) *Nature*, **328**, 496–500.
Sanger,F., Nicklen,S. and Coulson,A.R. (1977) *Proc. Natl Acad. Sci. USA*, **74**, 5463–5467.
Schechter,I. and Berger,A. (1967) *Biochem. Biophys. Res. Commun.*, **27**, 157–162.
Shaw,W.V. (1987) *Biochem. J.*, **246**, 1–17.
Teplyakov,A.V., Kuranova,I.P., Harutyunyan,E.H., Vainshtein,B.K., Frömmel,C., Höhne,W.E. and Wilson,K.S. (1990) *J. Mol. Biol.*, **214**, 261–279.
Tronrud,D.E., Ten Eyck,L. and Matthews,B.W. (1987) *Acta Crystallogr.*, **A43**, 489–501.
van Velzen,A.G. (1970) British Patent No. 1353317.
Wells,J.A., Powers,D.B., Bott,R.R., Graycar,T.P. and Estell,D.A. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 1219–1223.
Young,F.E., Smith,C. and Reilly,B.E. (1969) *J. Bacteriol.*, **98**, 1087–1097.

Received on October 31, 1991; revised and accepted on April 13, 1992